REMARKS/ARGUMENTS

The Status of the Claims.

Claims 21, 34, 35, 39, 41 to 44, 46 to 58 are pending with entry of this amendment. Claims 1 to 30, 32, 33, 36 to 38, 40, and 45 are cancelled. Claims 48 to 58 being added herein. Claims 31 and 47 are amended herein. These amendments introduce no new matter and support is replete throughout the specification. These amendments are made without prejudice and are not to be construed as abandonment of the previously claimed subject matter or agreement with any objection or rejection of record.

With respect to claims 31, support for the aspect of two or more different redox amino acids can be found throughout the specification. For example, see specification at paragraphs 4, 89 and 168.

With respect to claim 47, support for at least one redox amino acid in a protein can be found throughout the specification.

With regard to new claims 48 to 54, support for the specific structural aspects of the O-RSs can be found throughout the specification, e.g., at paragraphs 105 and 185.

Support for the 95% identity aspect in new claim 55 can be found, e.g., at paragraph 26.

Support for new claims 56 to 58 can be found throughout the specification. These aspects have been separated herein from independent claim 31 in the form of new dependent claims.

At page 2 of the Action, the Office identifies a discrepancy in the claim status of claims 36 and 37. The current claim listing acknowledges the cancelled status of these claims.

Applicants submit that no new matter has been added to the application by way of the above Amendment. Accordingly, entry of the Amendment is respectfully requested.

Claim Objections.

At page 3 of the Action, claim 37 is objected to for certain informalities. In the current claim set, claim 37 is cancelled. The objection should be withdrawn

35 U.S.C. §112, First Paragraph.

Claim 47 is rejected under 35 U.S.C. §112, first paragraph, for alleged lack of adequate written description and for alleged lack of enablement. To the extent the rejections are deemed applicable to the amended claims, Applicants traverse.

The written description requirement may be satisfied if claim terms "readily convey distinguishing information concerning their identity, such that one of ordinary skill in the art could visualize or recognize the identity of a member of the genus." See, *Amgen Inc. v. Hoescht Marion Roussel, Inc.* 65 USPQ2d 1385 (Fed. Cir. 2003). According to MPEP 2163, the description need only describe in detail that which is new or not conventional. The written description requirement for a claimed genus may be satisfied through sufficient description of a representative number of species by actual reduction to practice, reduction to drawings, or by disclosure of relevant, identifying characteristics, i.e., structure or other physical and/or chemical properties, by functional characteristics coupled with a known or disclosed correlation between function and structure, or by a combination of such identifying characteristics, sufficient to show the applicant was in possession of the claimed genus. See also, *Eli Lilly*, 119 F.3d at 1568, 43 USPQ2d at 1406. Here, the original specification allows one of skill to unarguably identify a member of the genus. Further, the original specification provides one of skill in the art methods to prepare compositions across the genus and provides structure/function information sufficient to show that the Applicants were in possession of the claimed genus, as discussed below. Therefore, written description of the invention is more than adequate, according to the law and examination guidelines.

To be an enabling disclosure under § 112, first paragraph, a patent must contain a description that enables one skilled in the art to make and use the claimed invention. That some experimentation is necessary does not constitute a lack of enablement; the amount of experimentation, however, must not be unduly extensive. See *In re Wands*, 8 USPQ2d 1400, 1404 (Fed. Cir. 1988). Whether undue experimentation is required by one skilled in the art is typically determined by reference to eight factors considered relevant to

the inquiry: (1) quantity of experimentation necessary; (2) amount of guidance presented; (3) presence of working examples; (4) nature of the invention; (5) state of the prior art; (6) relative skill of those in the art; (7) predictability of the art; and (8) breadth of the claims. See *id.*

Claim 47 is rejected based on the allegation that some of the variants of synthetase (O-RS) of SEQ ID NO.: 1 "do not necessarily have the function." In particular, the Office briefly points to the predictability of the art, working examples, and guidance factors of *Wands* analysis as allegedly failing in the instant disclosure. However, in light of the state of the art and given original disclosure, one of skill knows functional O-RS variants can be readily provided and identified across the genus of the claims.

The Office currently acknowledges enablement only of a past tense working example provided in the original specification, without acknowledging that one of skill can practice anything broader without undue experimentation. Such a stand is unreasonable, in light of the controlling *Wands* case where the Court found enablement of generic claims to monoclonal antibodies to an antigen where no structural information for the antibody was provided and experimentation routinely lead to a failure rate of 98% or higher. Here, Applicants have provided abundant structural information regarding all elements of the claims and guidance allowing a high degree of success in conservative substitutions of these components. Moreover, even without the benefit of the provided structure/functional information, the specification provides functional structures and powerful positive and negative selection techniques that have been shown to readily identify functional variants even from randomly mutated libraries.

The Examiner notes at page 7 of the Action that "[o]ne of skill in the art cannot predict a priori which polypeptides could be made 90% identical and have the requisite function, particularly when the consequences of substitution cannot be determined a priori." The statement can not form the basis for an enablement rejection. <u>Applicants note that consequences of experimentation do not have to be determined in advance (a priori), or they would literally not be experiments</u>. The basis of the rejection conflicts with logic and controlling law. However, the high skill in the art and guidance of the present specification provide abundant information on what peptide or nucleic acid residues of system components

can be conservatively substituted with a reasonable probability of retaining a useful degree of functionality, with routine experimentation. Applicants believe the original specification provides exhaustive structural information to guide one of skill in logical and predictable conservative sequence substitutions, likely to have a high degree of success. However, such enablement aside, one of skill is additionally enabled by the functional starting structures and screening methods provided in the specification that one of skill can undeniably use to readily obtain functional variants through routine experimentation. That is, Applicants have enabled one of skill to readily practice the compositions through at least two routes. The rejection must be withdrawn because 1) one of skill can rationally select variants with a high degree of predictability that function adequately as the claimed compositions, and 2) one of skill can start with the given sequences and use the given screening techniques to readily obtain functional conservative variants across the scope of the claimed compositions. Even if the specification lacked a discussion of structure-function relationships (which is does not lack) for direct engineering of compositions, screening methods could readily obtain functional variants throughout the genus starting from the provided functional structures.

The quantity of experimentation necessary to practice the claimed invention is the first *Wands* factor for evaluation of enablement. Even where some of the conservative variants generated by one of skill have reduced function, the *Wands* Court has found this to be quite acceptable. The Court first noted that the experimental process for making antibodies that bind the relevant antigen were set forth in the application (as here). In essence, this process included an elaborate hybridoma fusion screening and manipulation procedure, followed by a binding screen to identify "high binders" followed by another screening procedure to identify what type of antibody had been generated (IgM being the desirable antibody type in *Wands*). The PTO argued that less than 3% of hybridomas that were created produced antibodies, and of these, only 20% produced IgM antibodies. The first four hybridoma fusion experiments performed by the *Wands* inventors were failures, with the next 6 being successful. The Court held that this was not evidence of unpredictability, particularly given that the technique at issue was in general use for antibody production. *Wands* at 1406. Here, on the first try, Applicants have found a working system of the claim without the benefit of the teachings they now provide. Given the base structures

(not provided at all in *Wands*) one of skill can readily provide additional conservative variant species across the range of the claimed subject matter.

Here, a working reference sequence is given that will work every time, as compared to the *Wands* screening method that has to start from scratch and is known to fail a in substantial portion of attempts. Functional variants can be found more easily than in *Wands* because positions and functions of structures in tRNAs and synthetases, including conserved regions, binding pockets and active sites, are known. For example, tyrosyl tRNA synthetase is a Class I synthetase that contains a classic "Rossman Fold" nucleotide binding domain with a central 5-stranded parallel beta-sheet. See, e.g., *Rould et al.* (1989) Science 246, 1135-1142; *Brick et al.* (1988) J. Mol. Biol. 208, 83-98; *Biou et al.* (1994) Science 263, 1404-1410; *Perona et al.* (1993) Biochemistry 32, 8758-8771; and, *Delarue* (1995) Curr. Opin. Struct. Biol. 5, 48-55.

Applicants note that the best reference the Office could provide for the notion of unpredictable amino acid substitution was the ancient 1976 Rudinger reference (before the biotech revolution even got started - Applicants have been unable to obtain a copy of this reference and request one be provided by the Office). Things have changed in the field of protein engineering since then. The field is particularly mature in the field of translation systems. Indeed, it is likely that more is known about synthetase and tRNA structure than any other biomolecules, due to more than 50 years of intense study of there molecules. For example, the secondary and tertiary structures of thousands of synthetases are known. The folding structures of essentially all of the many thousands of known tRNAs are also known. Further, the current specification provides particular structural and functional information, including identification of key structures, such as identification of residues at positions corresponding to Tyr 32, Glu 107, Asp 158, Ile 159, and Leu 162 in the active site of MjTyrRS functioning to interact with moieties at the para position the aryl ring of tyrosine analogs, and residues at positions corresponding to Tyr32, Ala 67, His 70, Gln 155, Asp 158, Ala 167 in the active site functioning to interact with moieties at the meta position of the aryl ring of tyrosine and phenylalanine analogs. See, e.g., the original specification at starting at paragraph 185. Non-essential positions outside of the active site are known and casually modified with confidence as screening tools (see certain examples at paragraph 186). Further

structure and function information is available in Figures 4 to 6, and the sequence listing. With knowledge of the art and the information provided specifically guiding one of skill, a broader acknowledgement of enablement is clearly appropriate in this case.

Even at levels lower than the basic skill in the art, one with basic knowledge of general protein chemistry and genetics can identify functional conservative variations with a predictability well beyond the *Wands* standard. It goes without saying that one of skill could clearly practice the working example, acknowledged as enabled. However, something more must reasonably be considered enabled. For example, one of skill in the art of protein engineering, given the reference sequence and synthetase structural information in the specification and the public domain could surely succeed with minimal experimentation in logical directed conservative substitution of amino acids along the given sequence. For example, avoiding known and identified active site elements, one could easily substitute amino acids conducive to maintenance of scaffolding structures, such as α helices and β sheets, without disrupting secondary and tertiary conformation (Applicants note that controlling law, such as *Wands*, does allow for substantial rates of failure). Even if such substitutions influenced peptide activity, the result would often be a changed activity without total loss of function. In a pioneering reference "Progress toward the evolution of an organism with an expanded genetic code", by Liu, et al., PNAS 96: 4780-4785, at 4782, it is noted that orthogonal components charging 20x less than wild type is sufficient for function.

The Office has not provided any citation contradicting the assertions that logically directed substitutions are routine, and that even reduced activity still represents useful functionality. Modern references cited by the Office typically express amazement at finding amino acid substitutions in an active site that change activity of a peptide. These results are published in prestigious journals. It is hard to find the relatively rare mutations that destroy functionality, particularly outside of an active site (all of which are known for the synthetases of the invention). One of skill knows that anyone can readily identify functional conservative substitutions that do not change activity of a peptide, and that is why these observations of little interest and are rarely, if ever published.

One of skill in the art can readily prepare functional compositions of the claims, given the starting functional structures, without any consideration of any

structure/function relationships, known or unknown. Using the screening techniques provided in the specification, Applicants were able to provide the specific functioning constructs, without undue experimentation. Applicants provided the structures of the claimed methods starting from non-functional structures. Surely, it cannot be considered undue experimentation to accomplish the easier task of starting with functional structures and using the provided mutation and screening techniques to obtain additional functional variants.

In light of the above, experimentation would be minimal, and certainly not "undue" in this art, to successfully practice the present invention, e.g., by logical conservative directed substitution of any reasonable number amino acids in the reference O-RS, or even by random mutation of the given functional structures and screening for related functioning variants.

The amount of guidance presented is substantial, and well beyond that provided in *Wands*. The Action at page 9 suggests "no guidance regarding the residues that bear upon structure" are provided. However, this is not true. As discussed above, the original specification provides an entire working structure with known function of every component, e.g., including physical structure components and active site components. Structural members of synthetases "non-essential" to activity are well known in the art (the specification is not required to expound on what one of skill is expected to know). These structures make up the bulk of amino acids in the synthetase. It is known how to predictably substitute at these positions to retain secondary and tertiary structures that maintain active site coordination.

Alternate resources for system components are identified in the specification (see, paragraphs 85, and 87) as are conservative substitutions (see, paragraph 137). Also included are extensive citations of references (e.g., Anderson et al., (2002) Exploring the Limits of Codon and Anticodon Size, Chemistry and Biology, 9:237-244; GENBANK; computer-assisted modeling Macromodel version 8.1, Schrodinger, LLC; WO 2002/085923, "In Vivo Incorporation of Unnatural Amino Acids", by Shultz, et al., and WO 2002/086075, "Methods and Compositions for the Production of Orthogonal tRNA - Aminoacyl tRNA Synthetase Pairs" by Shultz, et al.) that would help guide one of skill to identify logical

conservative substitutions, within the limitations of the claims, retaining substantial functionality without undue experimentation.

Where a construct has previously been characterized, the construct can be transduced, transformed, or transfected into host cells for expression and production of the O-RS (and/or O-tRNA) of the invention. In many cases, a library of alternate candidate constructs is prepared, e.g., for a series of expression, screening, and selection steps to identify further constructs. The specification guides one of skill in steps to obtain functional system constituents from non-functional constituents by, e.g., preparing constructs of foreign RS/tRNA pairs, preparing libraries of mutant constituents, selecting for orthogonal constituents that do not that charge in the endogenous system, screening for suppressing pairs, screening for suppressing pairs that do not charge with natural amino acids, and screening for suppressing pairs that do charge with a redox amino acids. Given this guidance and the sequences provided, one of skill can practice the full scope of claimed methods with much less effort, e.g., skipping procedures already completed, and screening from libraries with a far higher percentage of functional members. For example, to practice composition variants, one of skill would not have to prepare constructs of foreign RS/tRNA pairs, not have to select for orthogonal constituents that do not that charge in the endogenous system, not have to screen for suppressing pairs, and/or not have to screen for suppressing pairs that do not charge with natural amino acids. Because the starting sequences already function, site directed mutant libraries (using structural information provided), or randomly mutated libraries, would necessarily include a far higher population of functional members. The percentage of suppressing pairs that charge with para or meta-substituted tryptophan would be orders of magnitude higher. Using the present functioning method and starting sequences, and the guidance of the specification, one of skill can practice the claimed methods with a much higher degree of success and predictability than the scientists of *Wands* or the work of the present inventors.

The techniques of positive and negative screening are capable of identifying, e.g., functional orthogonal synthetases charging unnatural amino acids, e.g., from libraries of randomly mutated synthetases; wherein the parent synthetase starts out totally non-functional in the desired orthogonal system. For example, see Liu, *ibid* at 4782, where individual

sessions of positive screening provided 130 to 10^5-fold enrichment for functional RSs, and sessions of negative screening provided 4000 to 10^7-fold enrichment for functional RSs. Here, Applicants have provided parent sequences (e.g., O-RS SEQ ID NO: 1), which already have functional structures. Using the techniques provided, one of skill can easily identify functional mutants based on the identified sequences. Continued rejection requires the Office to argue that where one of skill in the art can identify a functioning system by mutation of a non-functioning system (as in Lui), it is still undue experimentation to identify a functional system from mutants starting with a functional system. This is an objectively unreasonable argument.

The state of the art and the relative skill of those in the art are greater than in *Wands*. *Wands* indicated that the state of the prior art was advanced, with "all of the methods required to practice the invention being known." This is precisely true for the present case as well. Every step used to produce the claimed compositions is known and available, though some, such as the positive-negative screen combine several known methods to achieve the screen. Indeed, given that *Wands* was decided in 1988, it is plain that the state of the prior art is enormously more advanced than it was at the time of the *Wands* decision. The level of skill of practitioners in the field was considered "high" for the *Wands* decision. Obviously, it is much higher now than it was in 1988. The information that biotechnology practitioners are presumed to be aware of has had over 20 years to develop, and the pace of development during that period has been staggering. A typical postdoctoral researcher or principal investigator can, for example, sequence and provide a detailed analysis of an entire genome, or, e.g., hundreds of cloned RS or O-tRNA, in a matter of days or weeks, whereas in 1988, a week could go by to get one simple sequencing reaction to work, due to the extensive manual manipulations that had to be performed. If the level of skill in the art was "high" at the time of *Wands* then it is now positively stratospheric. In any case, any moderately competent molecular biologist, given Applicants' disclosure can certainly perform each and every step required to make the claimed compositions.

A very wide variety of tRNA synthetases were known at the time the invention was filed. For example, Szymanski et al. (2001) "Aminoacyl-tRNA synthetases database" Nucleic Acids Res. 29:288-290 (attached in Appendix A) provides one example

database of Aminoacyl-tRNA synthetases. See also: http://rose.man.poznan.pl/aars/index.html. Many synthetases have been described by sequence and are available in the literature—the one web site noted above lists over 1,000 available synthetases, including well over 100 for which the three dimensional crystal structure has been determined. This is one of the most studied and structurally characterized protein families in all of biology. One of skill knew functions of RS structures at the time, so could have practiced functional variants of a given structure without undue experimentation.

The predictability of the art is good, particularly here, where the general structure of a protein is given, including known binding pockets, structural scaffolding, and active sites. The rejection appears to be grounded on worst case scenarios that no mutated protein would be active or that one could not predict in advance what protein in a 10% mutated library would have activity. Applicants note that this is an improper interpretation of the law and guidelines. Yet, even with this worst case scenario, one of skill could review given sequences of the mutated active protein and make reasonable predictions as to which would be most likely to retain activity. However, the claims are all the more enabled because the artificial worst case scenario of the rejection fails to consider the contributions of the specification. <u>Experimentation does not require a predetermined result, only a reasonably predictable rate of success in experimentation.</u>

One of skill knows how to make conservative substitutions in less critical areas to retain the secondary and tertiary structures that retain the conformation of active sites to predictably provide a reasonable proportion of functioning proteins. For example, one of skill can have high confidence in substituting one amino acid known to cooperate in α helix stabilization (e.g., alanine) with another amino acid known to cooperate in α helix stabilization (e.g., leucine) with high confidence in retaining functionality of the peptide. This particularly when conserved structures, such as binding pockets and active sites have been identified (as in the present invention). Again, Applicants note that, even without systematic protection of conserved structures, and structures identified as to function, one of skill could predictably screen for functional variants of the provided sequences, even from a randomly mutated library. The predictable success rate would jump astronomically were the mutations logically directed in recognition of functional structures known in the art and

provided in the present specification. Enablement does not require absence of failure, accurate prediction of specific results beforehand, or optimum performance of each functional embodiment. In fact, failure was high in *Wands* and many antibodies had sub-optimal affinity for the antigen. Here, failure would be predictably low, and well below the controlling standard set in *Wands*. In fact, it would be unreasonable to argue that a range of functional variants could not be found with reasonable experimentation.

Because one of skill can readily practice desired species across the scope of the claims, alternately using logical engineering based on structures provided or by random mutation and screening of provided structures using provided techniques, Applicants respectfully request withdrawal of the rejections.

35 U.S.C. §103(a).

Claims 31, 34, 35, 38, 39, and 41 to 44 were rejected under 35 U.S.C. §103(a) as allegedly obvious based on Schultz (U.S. 7,045,337) in light of Rodriguez (Biochem. J. 149, 115-121 (1975)). To the extent the rejection is deemed applicable to the amended claims, Applicants traverse.

A proper analysis under the recently reaffirmed *Graham v John Deere* standard demonstrates the non-obviousness of the invention. According to the Supreme Court in *KSR International Co v. Teleflex* (550 U.S. ____ (2007); 127 S. Ct. 1727, 1740-41, 82 USPQ2d 1385-1396 (US 2007)), the appropriate standard for analyzing questions of obviousness is that:

> the scope and content of the prior art are determined, differences between the prior art and the claims at issue are analyzed and the level of ordinary skill in the pertinent art is resolved. Against this background the obviousness or non-obviousness of the subject matter is determined. Such secondary considerations as commercial success, long felt but unresolved needs, failure of others, etc. might be utilized to give light to the circumstances surrounding the origin of the subject matter to be patented.
>
> *Id.* quoting *Graham v. John Deere of Kansas City* 383 U.S. 1, 17-18.

The current Examination Guidelines (e.g., MPEP 2143) and *KSR* require the Office in an obviousness rejection to provide a statement as to why one of skill would have combined known elements. Further, an obviousness rejection must include fact-based

findings demonstrating: 1) a combination of reference elements describing each limitation of the claims, 2) known elements that function in the same way in the combination as in the references themselves, 3) the elements are combined by known methods, 4) the result of the suggested combination of elements would have been predictable, and 5) one of skill in the art would have expected success in providing the claim in light of the references. Here, the rejection fails each of these requirements, as applied to the *Graham* factors.

The Action alleges the existence of unnatural redox amino acids in the context of the orthogonal translation system components of Schultz '337. Rodriguez describes a rat brain soluble fraction that may inexplicably incorporate at most a single DHP to the C-terminus of a peptide. The Action does not state a finding of facts, required by MPEP 2143, explaining how the combination of elements would be obvious to one of skill based on the cited references.

The Office has not alleged all limitations of the claims are taught by the combination of references. In the final Office Action of February 7, 2008, the Office acknowledged that "Schultz et al do not specifically exemplify a myoglobin protein comprising at least two redox-active amino acids." With regard to the currently amended independent claim 31, the Office must provide findings of fact demonstrating that the combined references teach, e.g., 1) the combination of elements in the claim, and 2) a protein comprising at least two or more different redox amino acids selected from the claim list.

Even assuming all limitations were present in the combination of references, the elements would function differently and would not be combined by known methods. For example, Rodriguez does not know for sure what he has prepared (a "brain protein", see pages 116), nor does he know what mechanism appears to bind certain amino acids to what peptide ("soluble fraction of rat brain" somehow catalyzes C-terminal amino acid incorporation, see Discussion page 120). Therefore, Rodriguez does not provide known elements combined by known methods to reasonably support of any rejection. Further, the results of any combination with Rodriguez would predictably fail because he does not teach known substrates of reaction processes, but does confirm that his technique incorporates at most one amino acid to a peptide.

Applicants respectfully request withdrawal of rejections for alleged obviousness.

CONCLUSION

In view of the foregoing, Applicants believes all claims now pending in this application are in condition for allowance. The issuance of a formal Notice of Allowance at an early date is respectfully requested.

If the claims are deemed not to be in condition for allowance after consideration of this Response, a telephone interview with the Examiner is hereby requested. Please telephone the undersigned at (510) 769-3510 to schedule an interview.

Respectfully submitted,

Gary Baker

Gary Baker
Reg. No: 41,595

QUINE INTELLECTUAL PROPERTY LAW GROUP
P.O. BOX 458, Alameda, CA 94501
Tel: 510 769-3510
Fax: 510 337-7877
PTO Customer No.: **22798**
Deposit Account No.: **50-0893**
Attachments:

1) A transmittal sheet;
2) Appendix A - Szymanski et al. (2001), and,
3) A receipt indication postcard.

288–290 Nucleic Acids Research, 2001, Vol. 29, No. 1

Aminoacyl-tRNA synthetases database

Maciej Szymanski, Marzanna A. Deniziak and Jan Barciszewski*

Institute of Bioorganic Chemistry of the Polish Academy of Sciences, Noskowskiego 12/14, 61-704 Poznan, Poland

Received September 25, 2000; Revised and Accepted October 17, 2000

ABSTRACT

Aminoacyl-tRNA synthetases (AARSs) are at the center of the question of the origin of life. They constitute a family of enzymes integrating the two levels of cellular organization: nucleic acids and proteins. AARSs arose early in evolution and are believed to be a group of ancient proteins. They are responsible for attaching amino acid residues to their cognate tRNA molecules, which is the first step in the protein synthesis. The role they play in a living cell is essential for the precise deciphering of the genetic code. The analysis of AARSs evolutionary history was not possible for a long time due to a lack of a sufficiently large number of their amino acid sequences. The emerging picture of synthetases' evolution is a result of recent achievements in genomics [Woese,C., Olsen,G.J., Ibba,M. and Söll,D. (2000) *Microbiol. Mol. Biol. Rev.*, 64, 202–236]. In this paper we present a short introduction to the AARSs database. The updated database contains 1047 AARS primary structures from archaebacteria, eubacteria, mitochondria, chloroplasts and eukaryotic cells. It is the compilation of amino acid sequences of all AARSs known to date, which are available as separate entries via the WWW at http://biobases.ibch.poznan.pl/aars/.

INTRODUCTION

Most living cells possess a set of 20 aminoacyl-tRNA synthetases (AARSs), specifically charging their cognate tRNAs (1,2). Although some exceptions to the '20 AARSs' rule were found in bacteria, archaea and eukaryotic organelles (2), those were assumed to be just an evolutionary anomaly. However, recent discoveries arising mainly from functional genomics studies in bacterial and archaeal systems contradict these expectations. Numerous organisms do not possess a full complement of 20 AARSs. In many cases a reduced number of synthetases is supported by a variety of novel enzymes and pathways providing the full complement of aminoacyl-tRNAs required for protein synthesis (3).

The archaebacterial cysteinylation pathway is one of the recent discoveries in that field. It is known that the genome sequences of certain *Archaea* do not code for canonical cysteinyl-tRNA synthetases (CysRSs). The first proposed candidate for Cys-tRNA^{Cys} formation was seryl-tRNA synthetase. It was assumed that serine mischarged by SerRS could be potentially converted to cysteine. However, it has been found that SerRS from *Methanobacterium thermoautotrophicum* and *Methanococcus maripaludis* do not charge tRNA^{Cys} with serine (4). Surprisingly, cysteinyl-tRNA synthesis in *Methanococcus jannaschii* is mediated by the enzyme whose amino acid sequence is homologous to ProRS (5,6). Biochemical and genetic analyses indicated that the archaeal form of ProRS synthesizes both cognate cysteinyl-tRNA^{Cys} and prolyl-tRNA^{Pro}. This AARS of double specificity contains only one set of sequence motifs characteristic for class II enzymes. The whole protein is relatively small, ~53 kDa. Cloning of the enzyme in *Escherichia coli* shows that a single subunit supports both activities. Cross inhibition between specific proline and cysteine substrates also suggests a linkage or overlap between their active sites. It seems that the new archaeal enzyme is unlikely to be a multidomain peptide, in contrast to metazoan GluProRS. However, there is a controversy concerning cysteinylation and prolylation activity levels as well as tRNA^{Cys} requirement for activation of cysteine (5,6). It was suggested that the *M.jannaschii* ProRS might recruit additional protein or RNA factors to facilitate cysteinylation (6). Nevertheless, the ability of a single AARS to synthesize two aminoacyl-tRNAs raises questions about substrate specificity in protein synthesis and may provide insights into the evolutionary origin of the charging reaction (5–7). That kind of dual activity is observed not only in *Archaea*. The eukaryote *Giardia lamblia* also contains a ProRS with CysRS activity (2).

The recognition of a specific tRNA by its cognate synthetase depends on structural features of both molecules (8,9). There are numerous examples where tRNAs undergo large conformational changes upon binding to enzymes but little is known about conformational rearrangements in the tRNA-complexed AARSs. Recently the crystal structure of the dimeric class II aspartyl-tRNA synthetase (AspRS) from yeast was solved at 2.3 Å resolution. Its free form was compared with that of the protein associated with the cognate tRNA^{Asp} (10). Within the complex the AspRS undergoes some changes in the anticodon and catalytic domains of the enzyme subunits. The most significant differences are observed in the two evolutionary conserved loops, which are critical for ligand binding (10).

Interestingly certain AARSs, namely MetRS, IleRS and LeuRS, activate *in vivo* not only cognate amino acids but also the non-protein amino acid homocysteine (Hcy), due to its structural similarity to Met, Ile and Leu. Thus, Hcy enters the first steps of protein synthesis, but is never incorporated into the protein because of the editing mechanism of AARSs converting misactivated Hcy into thiolactone. Recently it was shown that reversibly modified Hcy (*S*-nitroso-Hcy) could bypass the editing step and be translationally incorporated into

*To whom correspondence should be addressed. Tel: +48 61 852 8503; Fax: +48 61 852 0532; Email: jbarcisz@ibch.poznan.pl

the protein by *E.coli* MetRS (11). This finding has some important medical implications, because the metabolic conversion of homocysteine to thiolactone and protein homocysteinylation may play a role in vascular damage (12).

The evolutionary history of AARS genes is difficult to establish due to numerous examples of gene duplications, fusions and horizontal transfers (13–15). Prokaryotes contain a set of 18–20 AARSs, but eukaryotes also possess their organellar analogs (16–18). Like mitochondria, plastids lost most of their original genes and some of them have been transferred to the nucleus (16). There are several plant AARSs for which one gene has been replaced by another one of different origin, but carrying the same function. There are two recently described AARSs among them: *Arabidopsis thaliana* asparaginyl-tRNA synthetase (AsnRS) and CysRS (19). Both cytosolic enzymes are very similar to their mitochondrial and plastidic counterparts, resulting in closely related AARSs in all three compartments. *Arabidopsis* genome analysis revealed the presence of four genes coding for AsnRSs and two sequences corresponding to CysRSs. Moreover, phylogenetic studies suggest that all of them have organellar origin and are thought to be the best examples known to date of the plastidic AARSs captured by the cytosolic protein synthesis machinery (19).

Multicellular eukaryotes contain a macromolecular assembly of nine AARSs and three auxiliary proteins. One of them, p43, is of special interest because it is a precursor of endothelial-monocyte-activating polypeptide II (EMAP II), an inflammatory cytokine involved in apoptosis (20). Recently the EMAP II portion of p43 has been localized within the multisynthetase complex isolated from rabbit reticulocytes (21). Immunoblot analysis has clearly demonstrated a strong reaction of anti-EMAP II antiserum with p43 as well as cross-reactivity with isoleucyl-tRNA synthetase. Electron microscopy images of the immunocomplexes show two antibody-binding sites: one is near the center of the multisynthetase complex, at the intersection of the arms with the base, the second is in the base of the particle. Those data allow a refinement of the three-domain model of polypeptide distribution within the multisynthetase complex. Moreover, the central location of the p43/EMAP II suggests that this polypeptide plays a role in optimizing the normal function and in a rapid disruption of the essential cellular machinery when apoptosis is signaled (21).

A great help in understanding diverse biological activities as well as the maturation process of EMAP II came from its crystal structure solved at 1.8 Å resolution (22). It revealed the tRNA-binding fold and a domain that is structurally homologous to other chemokines. The structures similar to the EMAP II RNA-binding motif were previously observed in the anticodon-binding domain of yeast AspRS and the B2 domain of *Thermus thermophilus* PheRS. The RNA-binding pattern of EMAP II is likely to be non-specific in contrast to the AspRS of *Saccharomyces cerevisiae*.

An assembly of AARSs complex seems to be mediated by the heat shock protein 90 (Hsp 90), which binds to human glutamyl-prolyl-tRNA synthetase (GluProRS) (23). This interaction is sensitive to the Hsp 90 inhibitor, geldanamycin and ATP. It is targeted to the region of three tandem repeats linking the two catalytic domains of GluProRS. The GluProRS interacts also with IleRS. This process depends on the activity of Hsp 90, implying that the association is mediated by the chaperone. It was found that Hsp 90 binds preferentially to most of the enzymes located in the complex, rather than to those that are not found there. In addition, inactivation of Hsp 90 interfered with the *in vivo* incorporation of the nascent AARSs into the multi-AARS complex. Thus, Hsp 90 appears to mediate protein–protein interactions of mammalian tRNA synthetases associating with their subset during complex formation (23). Concerning binding of Hsp 90 to GluProRS one should mention that the recently published NMR structure of one repeated motif from the linker part of the synthetase points to its possible interactions with RNA (24). This repeat is built around an antiparallel coiled-coil with the conserved lysine and arginine residues, which form a basic path on one side of the structure. This motif can be a docking site for nucleic acids. Gel retardation and filter-binding experiments confirmed the real interaction of the studied domain with RNA. Thus, it has been suggested that the GluProRS repeated motif might represent a novel type of general RNA-binding domain appended to eukaryotic AARSs, to serve as a *cis*-acting tRNA-binding cofactor (24).

The existence of AARSs in the nucleus was implied by a study on nuclear aminoacylation of tRNA (25–27). However, there was almost no information regarding the status of active AARSs within the nuclei of eukaryotic cells. Recently at least 13 active AARSs have been found in the purified nuclei of both CHO and rabbit kidney cells, although their steady-state levels represent only a small percentage of those found in the cytoplasm (28). Most interestingly, all the nuclear aminoacyl-tRNA synthetases examined so far can be isolated as art of a multienzyme complex that is more stable, and consequently larger, than the comparable complex isolated from the cytoplasm. These data directly demonstrate the presence of active AARSs in mammalian cell nuclei. Moreover, their unexpected structural organization raises questions concerning the functional significance of those multienzyme complexes and their role in nuclear to cytoplasmic transport of tRNAs. The role played by AARSs present in the nucleus is not limited to tRNA maturation and/or export control. Human methionyl-tRNA synthetase (MetRS) was shown to be translocated into the nucleolus in proliferating cells (29). This process depends on the integrity of RNA and the activity of RNA polymerase I in the nucleolus. The treatment of MetRS with specific anti-MetRS antibodies decreased ribosomal RNA synthesis. Thus, human MetRS is thought to play a role in the biogenesis of rRNA in nucleoli, while it is catalytically involved in protein synthesis in the cytoplasm.

There are many data showing that the functional units (modules) of AARSs are arranged along their sequence in a linear fashion. This seems to be a result of a long period of evolution, because numerous examples of synthetase-like proteins have been identified in recently sequenced genomes (30). Those proteins consist mostly of just one synthetase domain. A search of the current sequence databases enabled identification of alanyl-, aspartyl-, glutamyl-, glycyl-, histidyl-, lysyl-, methionyl-, phenylalanyl-, seryl- and tyrosyl-tRNA synthetase-like proteins (30). Some of them are limited to a single kingdom of the tree of life, whereas others are more widely distributed. Their genes are probably related to AARSs through gene duplication events, although details of these processes probably vary from one system to another (30).

DESCRIPTION AND AVAILABILITY OF THE DATABASE

The AARSs database is the collection of amino acid sequences of all published AARSs. Currently it contains 1047 primary structures of cytoplasmic and organellar AARSs from various organisms. The entries are grouped according to AARS amino acid specificity. They are based on EMBL/SWISS-PROT format. Each includes the AARS amino acid sequence, its SWISS-PROT name and the accession number, a short description of the sequence, its source (organism name with taxonomic classification) and bibliographic information. For the enzymes whose sequences were determined at the nucleotide level, the appropriate EMBL/GenBank or TIGR entries are included, and for those with already known 3D structure, the cross-references to the Brookhaven Protein Data Base are indicated. The partial sequences of AARSs are also included in the database. According to the original SWISS-PROT description, some of the entries have been marked as putative or probable. The AARSs database is available on the WWW at http://biobases.ibch.poznan.pl/aars/.

REFERENCES

1. Crick,F.H.C. (1958) On protein synthesis. *Symp. Soc. Exp. Biol.*, **12**, 138–163.
2. Ibba,M., Becker,H.D., Stathopoulos,C., Tumbula,D. and Söll,D. (2000) The Adaptor hypothesis revisited. *Trends Biochem. Sci.*, **25**, 311–316.
3. Woese,C., Olsen,G.J., Ibba,M. and Söll,D. (2000) Aminoacyl-tRNA synthetases, the genetic code, and the evolutionary process. *Microbiol. Mol. Biol. Rev.*, **64**, 202–236.
4. Kim,H.-S., Vothknecht,U.C., Hedderich,R., Celic,I. and Söll,D. (1998) Sequence divergence of seryl-tRNA synthetases in *Archaea*. *J. Bacteriol.*, **180**, 6446–6449.
5. Stathopoulos,C., Li,T., Longman,R., Vothknecht,U.C., Becker,H.D., Ibba,M. and Söll,D. (2000) One polypeptide with two aminoacyl-tRNA synthetase activities. *Science*, **287**, 479–482.
6. Lipman,R.S.A., Sowers,K.R. and Hou,Y.-M. (2000) Synthesis of cysteinyl-tRNA^{Cys} by a genome that lacks the normal cysteine-tRNA synthetase. *Biochemistry*, **39**, 7792–7798.
7. Yarus,M. (2000) Unraveling the riddle of ProCys-tRNA synthetase. *Science*, **287**, 440–441.
8. Beuning,P.J. and Musier-Forsyth,K. (1999) Transfer RNA recognition by aminoacyl-tRNA synthetases. *Biopolymers*, **52**, 1–28.
9. Sherlin,L.D., Bullock,T.L., Newberry,K.J., Lipman,R.S.A., Hou,Y.-M., Beijer,B., Sproat,B.S. and Perona,J. (2000) Influence of transfer RNA tertiary structure on aminoacylation efficiency by glutaminyl and cysteinyl-tRNA synthetases. *J. Mol. Biol.*, **299**, 431–446.
10. Sauter,C., Lorber,B., Cavarelli,J., Moras,D. and Giegé,R. (2000) The free yeast aspartyl-tRNA synthetase differs from the tRNA^{Asp}-complexed enzyme by structural changes in the catalytic site, hinge region, and anticodon-binding domain. *J. Mol. Biol.*, **299**, 1313–1324.
11. Jakubowski,H. (2000) Translational incorporation of S-nitrosohomocysteine into protein. *J. Biol. Chem.*, **275**, 21813–21816.
12. Jakubowski,H., Zhang,L., Bardeguez,A. and Aviv,A. (2000) Homocysteine thiolactone and protein homocysteinylation in human endothelial cells. *Circ. Res.*, **87**, 45–51.
13. Brown,J.R. and Doolittle,W.F. (1995) Root of the universal tree of life based on ancient aminoacyl-tRNA synthetase gene duplications. *Proc. Natl Acad. Sci. USA*, **92**, 2441–2445.
14. Diaz-Lazcoz,Y., Aude,J.-C., Nitschké,P., Chiapello,H., Landès-Devauchelle,C. and Risler,J.-R. (1998) Evolution of genes, evolution of species: the case of aminoacyl-tRNA synthetases. *Mol. Biol. Evol.*, **15**, 1548–1561.
15. Wolf,Y.I., Aravind,L., Grishin,N.V. and Koonin,E.V. (1999) Evolution of aminoacyl-tRNA synthetases – analyses of unique domain architectures and phylogenetic trees reveals a complex history of horizontal gene transfer events. *Genome Res.*, **9**, 689–710.
16. Mireau,H., Lancelin,D. and Small,I.D. (1996) The same *Arabidopsis* gene encodes both cytosolic and mitochondrial alanyl-tRNA synthetases. *Plant Cell*, **8**, 1027–1039.
17. Akashi,K., Grandjean,O. and Small,I. (1998) Potential dual targeting of an *Arabidopsis* archaebacterial-like histidyl-tRNA synthetase to mitochondria and chloroplasts. *FEBS Lett.*, **431**, 39–44.
18. Menand,B., Marechal-Drouard,L., Sakamoto,W., Dietrich,A. and Wintz,H. (1998) A single gene of chloroplast origin codes for mitochondrial and chloroplastic methionyl-tRNA synthetase in *Arabidopsis thaliana*. *Proc. Natl Acad. Sci. USA*, **95**, 11014–11019.
19. Peeters,N.M., Chapron,A., Giritch,A., Grandjean,O., Lancelin,D., Lhomme,T., Vivrel,A. and Small,I. (2000) Duplication and quadruplication of *Arabidopsis thaliana* cysteinyl- and asparaginyl-tRNA synthetase genes of organellar origin. *J. Mol. Evol.*, **50**, 413–423.
20. Quevillon,S., Agou,F., Robinson,J.-C. and Mirande,M. (1997) The p43 component of the mammalian multi-synthetase complex is likely to be the precursor of the endothelial-monocyte-activating polypeptide II cytokine. *J. Biol. Chem.*, **272**, 32573–32579.
21. Norcum,M.T. and Warrington,A. (2000) The cytokine portion of p43 occupies a central position within the eykaryotic multisynthetase complex. *J. Biol. Chem.*, **275**, 17921–17924.
22. Kim,Y., Shin,J., Li,R., Cheong,C., Kim,K. and Kim,S. (2000) A novel anti-tumor cytokine contains an RNA binding motif present in aminoacyl-tRNA synthetases. *J. Biol. Chem.*, **275**, 27062–27068.
23. Kang,J., Kim,T., Ko,Y.-G., Rho,S.B., Park,S.G., Kim,M.J., Kwon,H.J. and Kim,S. (2000) Heat shock protein 90 mediates protein-protein interactions between human aminoacyl-tRNA synthetases. *J. Biol. Chem.*, **275**, 31682–31688.
24. Cahuzac,B., Berthonneau,E., Birlirakis,N., Guittet,E. and Mirande,M. (2000) A recurrent RNA-binding domain is appended to eukaryotic aminoacyl-tRNA synthetases. *EMBO J.*, **19**, 445–452.
25. Hooper,A.K. (1998) Nuclear function charge ahead. *Science*, **282**, 2003–2004.
26. Lund,E. and Dahlberg,J.E. (1998) Proofreading and aminoacylation of tRNAs before export from the nucleus. *Science*, **282**, 2082–2085.
27. Sarkar,S., Azad,A.K. and Hooper,A.K. (1999) Nuclear tRNA aminoacylation and its role in nuclear export of endogenous tRNAs in *Saccharomyces cerevisiae*. *Proc. Natl Acad. Sci. USA*, **96**, 14366–14371.
28. Nathanson,L. and Deutscher,M.P. (2000) Active aminoacyl-tRNA synthetases are present in nuclei as high molecular weight multienzyme complex. *J. Biol. Chem.*, **275**, 31559–31562.
29. Ko,Y.-G., Kang,Y.-S., Kim,E.-K., Park,S.G. and Kim,S. (2000) Nucleolar localization of human methionyl-tRNA synthetase and its role in ribosomal RNA synthesis. *J. Cell Biol.*, **149**, 567–574.
30. Schimmel,P. and de Pouplana,L.R. (2000) Footprints of aminoacyl-tRNA synthetases are everywhere. *Trends Biochem. Sci.*, **25**, 207–209.